

(51) Int.Cl. ^a	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00			G 1 0 L 3/00	H
// G 0 6 F 17/22			G 0 6 F 15/20	5 2 0 L

審査請求 未請求 請求項の数21 O L (全 15 頁)

(21) 出願番号 特願平8-45125

(22) 出願日 平成8年(1996)3月1日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 大山 芳史

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 浅野 久子

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 松岡 浩司

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

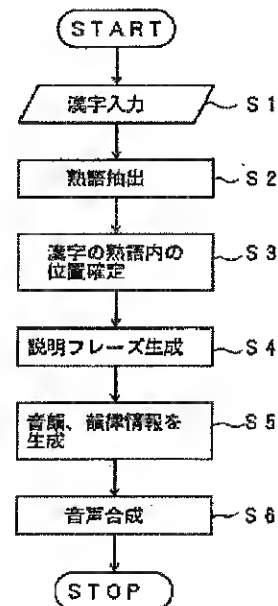
(54) 【発明の名称】 漢字説明方法及び装置

(57) 【要約】

【課題】 説明対象の漢字を個々に説明する文を、自動的に生成でき、しかも正しい読み情報とアクセント情報が付与できるため、自然な合成音声の出力が可能な漢字説明方法及び装置を提供することを目的とする。

【解決手段】 本発明は、入力された説明対象の漢字を含む熟語を辞書を参照することにより取得し、熟語の読み情報に基づいて前記漢字を説明するための熟語を選択し、正しい読み情報及び自立語の有無に応じて説明フレーズを生成し、該説明フレーズに音韻情報及び韻律情報を付与し、説明フレーズを音声合成して出力する。

本発明の原理を説明するための図



1

【特許請求の範囲】

【請求項1】 説明対象の漢字を個々に説明する文を生成して出力する漢字説明方法において、
入力された説明対象の漢字を含む熟語を辞書を参照することにより取得し、
前記熟語の読み情報に基づいて前記漢字を説明するための熟語を選択し、
正しい読み情報及び自立語の有無に応じて説明フレーズを生成し、該説明フレーズに音韻情報及び韻律情報を付与し、
前記説明フレーズを音声合成して出力することを特徴とする漢字説明方法。

【請求項2】 説明対象の漢字を受け付け、
各用語に対応した複数の表記、読み、品詞、アクセント情報を収録した辞書を用いて前記説明対象の漢字を含む熟語を抽出し、
抽出された前記熟語の表記の中で、入力された前記説明対象の漢字の位置を確定し、
該当する位置の漢字に対応する読みで代表される音韻情報及びアクセントに代表される韻律情報を生成して前記説明フレーズとし、
前記説明フレーズの前記音韻情報及び前記韻律情報に基づいて音声合成を行い、合成音声により前記漢字の説明を出力する請求項1記載の漢字説明方法。

【請求項3】 前記説明対象の漢字を含む熟語を抽出する際に、
音韻、又は、音韻と韻律が同じ、または、音韻と韻律が類似した熟語による多義の数の少ない方を選択する請求項2記載の漢字説明方法。

【請求項4】 前記説明対象の漢字を含む熟語を抽出する際に、
前記説明対象の漢字を含む熟語を複数抽出し、
抽出された熟語の漢字に対する説明が複数生成された場合に、熟語毎に該説明を併記して出力する請求項2記載の漢字説明方法。

【請求項5】 前記説明対象の漢字を含む熟語を抽出する際に、
入力された前記漢字をキーにして前記漢字の表記、読み、アクセント情報を保持する熟語辞書を検索して、前記漢字を含む熟語を抽出し、
前記漢字の表記、読み、品詞、アクセント情報を保持する音訓辞書を参照して前記熟語の表記に対する入力された前記漢字の位置と読みの候補を取得する請求項2記載の漢字説明方法。

【請求項6】 前記説明対象の漢字を含む熟語を抽出する際に、
それまでに既に抽出されている熟語を除外し、他の熟語を選択する請求項2記載の漢字説明方法。

【請求項7】 説明対象の漢字を個々に説明する文を生成して出力する漢字説明方法において、

2

漢字を形状を組み合わせから構成される辞書を用いて、
入力された説明対象漢字の説明フレーズを生成し、
前記説明フレーズを音声合成することにより出力することを特徴とする漢字説明方法。

【請求項8】 前記説明フレーズを生成する場合に、前記辞書の漢字に対応する説明情報と該漢字に関する構成情報を参照する請求項7記載の漢字説明方法。

【請求項9】 説明対象の漢字を入力する漢字入力手段と、

10 前記漢字入力手段により入力された前記漢字に対応する表記、読み、品詞、アクセント情報を保持する辞書と、
前記説明対象の漢字を含む熟語を、前記辞書を検索して抽出する熟語抽出手段と、
前記熟語抽出手段により抽出された熟語の表記より前記漢字の位置を確定する位置確定手段と、
前記熟語の表記と前記位置確定手段により確定された前記漢字の位置に対応する音韻情報及び韻律情報を用いて説明文を生成する説明文生成手段と、
前記説明文生成手段により生成された説明文を音声により出力する音声出力手段とを有することを特徴とする漢字説明装置。

【請求項10】 前記熟語抽出手段は、
音韻または、音韻と韻律が同じ、または、該音韻と韻律が類似した熟語のうち、多義の数の少ない方を選択する手段を含む請求項9記載の漢字説明装置。

【請求項11】 前記説明文生成手段は、
前記熟語抽出手段により、前記説明対象の漢字を含む熟語が複数抽出された場合には、前記説明対象の漢字の読みと熟語毎に説明文を併記して出力する手段を含む請求項9記載の漢字説明装置。

【請求項12】 前記辞書は、
漢字の表記、読み、アクセント情報を保持する熟語辞書と、
漢字毎の読み、品詞、アクセント情報を保持する音訓辞書とを含む請求項9記載の漢字説明装置。

【請求項13】 前記説明文生成手段は、
前記熟語抽出手段により、それまでに既に抽出されている熟語を除外し、他の熟語を選択して説明文を生成する熟語除外手段を含む請求項9記載の漢字説明装置。

40 【請求項14】 前記説明文生成手段は、
前記位置確定手段による前記説明対象の漢字の複数の熟語の各々の位置情報から該漢字の読みを選択する場合には、連濁などで生じる濁音、半濁音が含まれている場合には、清音化して一致する読みを採用する読み選択手段を含む請求項9記載の漢字説明装置。

【請求項15】 前記説明文生成手段は、
前記説明対象の漢字を含む熟語の読みの後に、助詞を挿入し、該助詞の後に前記漢字の読みを設定する説明文結合手段を含む請求項9及び14記載の漢字説明装置。

50 【請求項16】 前記説明文生成手段は、

前記説明対象の漢字の読みが自立語を表す場合には、熟語の読みと当該漢字の読みを助詞で接続した説明フレーズを作成し、自立語を表さない場合には熟語の読みと熟語の読みで当該漢字の読みの部分を助詞で接続した説明フレーズを作成する手段を含む請求項9及び15記載の漢字説明装置。

【請求項17】 前記漢字入力手段は、手書き文字を認識する文字認識手段を含む請求項9記載の漢字説明装置。

【請求項18】 前記音声出力手段は、前記音韻情報と前記韻律情報に基づいて音声を作成し、合成音声を出力する音声合成手段を含む請求項9記載の漢字説明装置。

【請求項19】 前記熟語辞書は、利用者の検索の指向性に応じた語彙を有する内容とする請求項12記載の漢字説明装置。

【請求項20】 説明対象の漢字を入力する漢字入力手段と、前記漢字入力手段により入力された漢字に対応する漢字の形状を組み合わせを有する漢字構成辞書と、前記漢字構成辞書に対応する音韻情報を格納する音韻辞書と、前記漢字構成辞書を検索して、前記説明対象漢字に対応する漢字の形状の説明を取得し、前記音韻辞書を参照して説明フレーズを生成する説明フレーズ生成手段と、前記説明フレーズ生成手段により生成された説明フレーズを音声合成して出力する音声合成手段とを有することを特徴とする漢字説明装置。

【請求項21】 請求項9及び請求項20記載の漢字説明装置を連動して利用する漢字説明装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、漢字説明方法及び装置に係り、特に、音声により漢字の表記を説明する情報案内、または、文字入力時の確認のための漢字情報案内を行う際の漢字説明方法及び装置に関する。

【0002】

【従来の技術】音声による同音語の確認には、『市立』と『私立』を区別するために、『私立』を「ワタクシリツ」、『市立』を「イチリツ」と読み上げる工夫は、ラジオ等のアナウンスにより工夫されている。

【0003】また、新聞の構成システムとして『追求』『ツイキュウ、モトム』のように予め特定の同音語に特別の呼びを付与することにより、区別するように読み上げるシステムが存在する。このようなシステムは、“日本文訂正支援システム(REVISE)”(日本電信電話株式会社研究実用化報告第336号第9号(1987))に開示されている。

【0004】

【発明が解決しようとする課題】しかしながら、従来の

技術における前者の方法は、システム化されておらず、後者は、校正の目的のために特別な読み方が定義されているため、一般の電話サービス等の広範囲な利用者を想定した場合、適用できないという問題がある。

【0005】また、個々の漢字の説明を同音語の有無に関わらず、自動的に読み上げて確認する方法は実現されていない。本発明は、上記の点に鑑みなされたもので、上記従来の問題点を解決し、説明対象の漢字を個々に説明する文を、自動的に生成でき、しかも正しい読み情報とアクセント情報が付与できるため、自然な合成音声の出力が可能な漢字説明方法及び装置を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明の漢字説明方法は、説明対象の漢字を個々に説明する文を生成して出力する漢字説明方法において、入力された説明対象の漢字を含む熟語を辞書を参照することにより取得し、熟語の読み情報に基づいて前記漢字を説明するための熟語を選択し、正しい読み情報及び自立語の有無に応じて説明フレーズを生成し、該説明フレーズに音韻情報及び韻律情報を付与し、説明フレーズを音声合成して出力する。

【0007】図1は、本発明の原理を説明するための図である。本発明の漢字説明方法は、説明対象の漢字を受け付け(ステップ1)、各用語に対応した複数の表記、読み、品詞、アクセント情報を収録した辞書を用いて説明対象の漢字を含む熟語を抽出し(ステップ2)、抽出された熟語の表記の中で、入力された説明対象の漢字の位置を確定し(ステップ3)、確定された漢字の位置に基づいて漢字の読み及び熟語の読みを取得して、漢字の説明フレーズを生成し(ステップ4)、該当する位置の漢字に対応する読みで代表される音韻情報及びアクセントに代表される韻律情報を説明フレーズに対応させて生成し(ステップ5)、説明フレーズの音韻情報及び韻律情報に基づいて音声合成を行い(ステップ5)、合成音声により漢字の説明を出力する(ステップ6)。

【0008】上記のステップ2において、説明対象の漢字を含む熟語を抽出する際に、音韻、又は、音韻と韻律が同じ、または、音韻と韻律が類似した熟語による多義の数の少ない方を選択する。上記のステップ2において、説明対象の漢字を含む熟語を抽出する際に、説明対象の漢字を含む熟語を複数抽出し、抽出された熟語の漢字に対する説明が複数生成された場合に、熟語毎に該説明を併記して出力する。

【0009】上記ステップ2において、説明対象の漢字を含む熟語を抽出する際に、入力された漢字をキーにして漢字の表記、読み、アクセント情報を保持する熟語辞書を検索して、漢字を含む熟語を抽出し、漢字の表記、読み、品詞、アクセント情報を保持する音訓辞書を参照して熟語の表記に対する入力された漢字の位置と読みの候補を取得する。

【0010】上記ステップ4において、熟語に説明が付与された情報を出力する際に、それまでに既に抽出されている熟語を除外し、他の熟語を選択する。また、本発明は、説明対象の漢字を個々に説明する文を生成して出力する漢字説明方法において、漢字を形状を組み合わせから校正される辞書を用いて、入力された説明対象漢字の説明フレーズを生成し、説明フレーズを音声合成することにより出力する。

【0011】また、本発明は、説明フレーズを生成する場合に、辞書の漢字に対応する説明情報と該漢字に関する構成情報を参照する。図2は、本発明の原理構成図である。本発明の漢字説明装置は、説明対象の漢字を入力する漢字入力手段10と、漢字入力手段10により入力された漢字に対応する表記、読み、品詞、アクセント情報を保持する辞書20と、説明対象の漢字を含む熟語を、辞書20を検索して抽出する熟語抽出手段30と、熟語抽出手段30により抽出された熟語の表記より漢字の位置を確定する位置確定手段40と、位置確定手段40により確定された漢字の位置に対応する音韻情報及び韻律情報を用いて説明文を生成する説明文生成手段50と、説明文生成手段50により生成された説明文を音声により出力する音声出力手段60とを有する。

【0012】上記の熟語抽出手段30は、音韻または、音韻と韻律が同じ、または、該音韻と韻律が類似した熟語のうち、多義の数の少ない方を選択する手段を含む。上記の説明文生成手段50は、熟語抽出手段30により、説明対象の漢字を含む熟語が複数抽出された場合には、説明対象の漢字の読みに熟語毎に各々説明文を併記して出力する手段を含む。

【0013】上記の辞書20は、漢字の表記、読み、アクセント情報を保持する熟語辞書と、漢字毎の読み、品詞、アクセント情報を保持する音訓辞書とを含む。上記の説明文生成手段50は、熟語抽出手段30により、複数の熟語が抽出された場合に、それまでに既に抽出されている熟語を除外し、他の熟語を選択して説明文を生成する熟語除外手段を含む。

【0014】上記の説明文生成手段50は、位置確定手段40による説明対象の漢字の複数の熟語の各々の位置情報から該漢字の読みを選択する場合に、連濁などで生じる濁音、半濁音が含まれている場合には、清音化して一致する読みを採用する読み選択手段を含む。

【0015】上記の説明文生成手段50は、説明対象の漢字を含む熟語の読みの後に、助詞を挿入し、該助詞の後に漢字の読みを設定する説明文結合手段を含む。上記の説明文生成手段50は、説明対象の漢字の読みが自立語を表す場合には、熟語の読みと当該漢字の読みを助詞で接続した説明フレーズを作成し、自立語を表さない場合には熟語の読みと熟語の読みで当該漢字の読みの部分を助詞で接続した説明フレーズを作成する手段を含む。

【0016】上記の熟語辞書は、利用者の検索の指向性

に応じた語彙を有する内容とする。上記の漢字入力手段10は、手書き文字を認識する文字認識手段を含む。上記の音声出力手段60は、音韻情報と韻律情報に基づいて音声合成し、合成音声を出力する音声合成手段を含む。

【0017】また、本発明の漢字説明装置は、説明対象の漢字を入力する漢字入力手段と、漢字入力手段により入力された漢字に対応する漢字の形状を組み合わせを有する漢字構成辞書と、漢字構成辞書に対応する音韻情報を格納する音韻辞書と、漢字構成辞書を検索して、説明対象漢字に対応する漢字の形状の説明を取得し、音韻辞書を参照して説明フレーズを生成する説明フレーズ生成手段と、説明フレーズ生成手段により生成された説明フレーズを音声合成して出力する音声合成手段とを有する。

【0018】さらに、本発明は、熟語の読みと説明対象漢字の読みを説明フレーズとして出力する装置と、漢字の形状を説明フレーズとして出力する装置を1つの装置として構成することも可能である。このように、本発明によれば、熟語等の辞書を準備することにより、同音語の有無に関わらず、漢字の確認のための文を必要に応じて音声で生成することが可能であり、電話等を解したサービスにおいて、騒音等で聞き取り難い場合は、区別の付きにくい類似した音韻の漢字を別の表現で漢字を説明しながら、情報を提供できるため、広範囲な確認の手段を提供することが可能である。

【0019】また、熟語中の説明対象漢字が自立語である場合には、当該漢字の読みを説明フレーズにそのまま挿入し、自立語でない場合には、熟語の読みの中の漢字の読みを挿入することにより、自立語でなくとも容易に当該漢字を認識することが可能となる。

【0020】また、本発明では、音韻情報のみならず、韻律情報を用いることにより、同じ読みであってもアクセントの位置の違いにより異なる文字を的確に区別することが可能である。また、熟語の辞書は、個人やグループに合った辞書を用意することで、その個人やグループにおいてよく使われる語彙を用いた説明も可能であるため、個人適応の漢字の説明を行うことも可能となる。

【0021】さらに、本発明は、漢字を構成する部首等の形状を組み合わせた説明を利用することが可能である。また、熟語の読み及び説明対象漢字の読みを説明フレーズとして出力し、更に、当該説明フレーズに加えて漢字を構成する部首等の形状の説明を行う説明フレーズを構成することも可能であり、画数の多い漢字や、適当な熟語がない漢字の説明を容易に行うことが可能となる。

【0022】

【発明の実施の形態】図3は、本発明の漢字説明装置の構成を示す。同図に示す漢字説明装置は、漢字入力部1、説明生成部2、音声合成装置3及び熟語辞書41と

音訓辞書42を有する用語辞書4から構成される。

【0023】漢字入力部1は、説明対象の漢字を入力する。入力された漢字が手書き文字の場合には、文字認識処理を行い、認識結果を漢字コードに変換して説明生成部2に転送する。説明生成部2は、図4に示すように、漢字入力部1において入力された説明対象の漢字を含む述語を用語辞書4から抽出する熟語抽出部21、熟語抽出部21で抽出された熟語の表記の中で、入力された漢字の位置を確定する位置確定部22、熟語抽出部21で抽出された熟語、当該熟語に含まれる説明対象漢字の読みに基づいて説明フレーズを生成する。また、熟語抽出部21により抽出された熟語が複数ある場合に、位置確定された説明対象漢字の読みにて、読みの連濁などで生じる濁音や半濁音を清音化して読みが一致する漢字を選択し、当該漢字の説明フレーズを生成する説明フレーズ生成部23、位置確定部22において確定された位置の漢字に対応する読みに代表される音韻情報、アクセントに代表される韻律情報とを説明フレーズに付与する音韻情報生成部24より構成される。

【0024】説明フレーズ生成部23は、熟語抽出部21により熟語辞書41が参照されて説明対象の漢字を含む熟語が検索されると、検索された熟語の表記に対する説明対象漢字の位置と、当該漢字の読みの候補を抽出する。説明対象の漢字を含む熟語を抽出する際に、音韻、または音韻と韻律が同じ、または、音韻と韻律が類似した熟語による多義の数が少ない方を選択する。また、熟語抽出部21から、説明対象の漢字を含む熟語が用語辞書4の熟語辞書41から複数抽出されると、当該熟語毎に、漢字に対する説明フレーズを生成し、さらに、音訓辞書42から抽出された漢字の表記に対する読み、品

【0025】次に、本発明の動作を説明する。図5は、本発明の漢字説明処理の一連の動作のフローチャートである。

ステップ100) 漢字入力部1は、説明対象漢字Kを入力する。この入力方法は、キーボード等から入力してもよいし、手書き文字をペン入力または、スキャナ等に読み込ませ、文字認識を行ってもよい。

【0026】ステップ110) 説明対象漢字Kを含む熟語を熟語辞書41を当該漢字Kをキーにして検索することにより熟語を取得する。

ステップ120) 取得した熟語がそれまでの対話等により用いた熟語である場合には、ステップ130に移行する。

【0027】ステップ130) 取得した熟語がそれまでの対話等により用いた熟語を除外する。除外した結果、熟語の数がもし0となった場合には、処理を終了する。

ステップ140) 説明生成部2は、熟語の選択処理を

行う。選択処理は、 n 個の熟語 $J(n)$ 毎に同音語の数 $D(J)$ を音訓辞書42を検索することにより取得し、当該同音語の数 $D(J)$ の数が最も少ない熟語を1つ選択して当該熟語を採用する方法と、取得した熟語の数 n が所定の数より少なければ、前述の処理を行わずに、ステップ150に移行するような方法等がある。

【0028】ステップ150) 説明生成部2は、単漢字Kをキーとして、音韻情報及び韻律情報を音訓辞書42から検索する。

ステップ160) 説明生成部2は、取得した音韻情報、韻律情報から以下のような構成の説明フレーズを生成する。

【0029】「熟語の読み」+「(の助詞)」+「Kの読み」

ステップ170) 説明生成部2は、生成された説明フレーズと熟語の数が一致すればステップ180に移行し、複数の熟語が抽出されている場合には、ステップ160の処理に移行し、説明フレーズを生成する。

【0030】ステップ180) 音声合成部3は、音韻情報、韻律情報に基づいて音声合成する。

ステップ190) 音声合成部3により合成された音声をスピーカ、電話機等より出力する。

【0031】次に、上記の説明生成部2の処理を詳細に説明する。図6は、本発明の説明生成部の処理を示すフローチャートである。

ステップ200) 漢字を説明するための熟語 $J(n)$ の候補 n は辞書上に複数存在するかを判定する。存在する場合には、ステップ210に移行し、存在しない場合には、ステップ220に移行する。

【0032】ステップ210) 複数の熟語 $J(n)$ が抽出された場合には、各熟語 $J(n)$ 毎にその熟語と音韻情報、韻律情報が同じまたは、類似している他の熟語に存在している数を調べ、最も少ない熟語 $J(n)$ を選択して $J(1)$ とする。

ステップ220) 入力された説明対象の漢字Kにて音訓辞書42を検索し、当該単漢字Kの読み等の情報を $K(s)$ を s 個抽出する。つまり、漢字Kに関する読みの候補を抽出する。

【0033】ステップ230) 位置確定部22は、選択された熟語 $J(1)$ の表記の単漢字Kの位置(左、中、右)を確定する。

ステップ240) 説明フレーズ生成部23は、 $K(s)$ による「読み」と熟語候補の漢字の位置(左、中、右)と単漢字の読みとして開始可能な条件から推定される単漢字 $K(s)$ の読みの連濁等で生じる、濁音や半濁音を除いて一致するものを抽出し、この単漢字を $K(1)$ とする。

【0034】ステップ250) 説明フレーズ生成部23は、単漢字 $K(1)$ は自立語であるかを判定し、自立語である場合には、ステップ260に移行し、自立語で

はない場合には、ステップ270に移行する。

ステップ260) 説明フレーズ生成部23は、単漢字K(1)が自立語である場合には、説明フレーズとして、

「J(1)の読み」+「の」+「K(1)の読み」を生成する。「J(1)の読み」は、熟語全体の読みを示し、「の」は接続するための助詞であり、「K(1)の読み」は、ステップ240で取得した単漢字Kの読みである。

【0035】ステップ270) 説明フレーズ生成部23は、単漢字K(1)が自立語でない場合には、説明フレーズとして

「J(1)の読み」+「の」+「J(1)中のK(1)の読み」

を生成する。単漢字が自立語でない場合には、最後に接続する読みは、抽出された熟語J(1)内に含まれる単漢字K(1)の読みを助詞(「の」)により接続する。

【0036】ステップ280) 韻律情報生成部24は、上記の説明フレーズに用語辞書4から取得したアクセント型に応じて音韻情報を付与する。次に、漢字を組み合わせて入力された説明対象漢字を説明する方法について説明する。

【0037】図7は、本発明の第2の漢字説明装置の構成を示す。同図に示す漢字説明装置は、漢字入力部5、説明フレーズ生成部6、音声合成部7、漢字形状辞書8及び音韻辞書9より構成される。漢字入力部5は、説明対象の漢字を入力する。入力された漢字が手書き文字の場合には、文字認識処理を行い、認識結果を漢字コードに変換して説明生成部2に転送する。

【0038】説明フレーズ生成部6は、漢字形状辞書8を参照して、入力された説明対象漢字の部首等の形状を説明する内容を取得し、さらに音韻辞書9を参照して、形状に関する説明の読みを取得して、説明フレーズを生成する。音声合成部7は、説明フレーズ生成部6により生成された説明フレーズを音韻情報に基づいて音声合成して、音声合成された説明文を出力する。

【0039】図8は、本発明の第2の漢字説明処理の動作のフローチャートである。

ステップ310) 漢字入力部1は、説明対象漢字を入力する。

ステップ320) 入力された説明対象漢字をキーとして漢字形状辞書8を参照して、音韻辞書を検索するための番号または、記号を取得する。

【0040】ステップ330) 漢字形状辞書8により取得した番号または記号を用いて音韻辞書9を検索し、当該漢字の形状に対応する音韻情報を取得する。

ステップ340) 音韻情報に基づいて説明フレーズを生成する。

ステップ350) 生成された説明フレーズに基づいて音声合成を行い、出力する。

【0041】

【実施例】以下、図面と共に、本発明の実施例を説明する。

【第1の実施例】図9は、本発明の実施例を説明するための図である。同図において、“ ”は、アクセントの上がり目を示し、“/”はアクセント句の境界を示す。同図は、説明対象漢字、使用する熟語、生成される説明、生成される説明のアクセント付きカナ名の各項目から構成される。

【0042】図10は、本発明の実施例の熟語辞書の例を示し、キー項目、見出し、読み、アクセント型等の項目より構成される。図11は、本発明の実施例の音訓辞書の例を示し、見出し及びキー、読み、品詞、アクセント型等の項目より構成される。

【0043】本実施例では、漢字入力部1から説明対象の漢字として『垣』が入力された場合について説明する。説明生成部2の熟語抽出部21は、『垣』をキーとして熟語辞書41を検索し、表記に対する説明対象漢字の位置とその漢字の読みの候補を抽出する。この例において、熟語抽出部21により抽出された熟語が「垣根」である場合、読みとして「カキネ」が取得され、アクセント型として“2”が取得される。

【0044】位置確定部22は、「垣根」の『垣』は、“根”の左側に位置すると判定する。これにより、読みの「カキネ」に対しては、左側の漢字「垣」に対する読みとしては、この段階において、「カキ」または「カ」が想定される。また、説明対象の漢字「垣」の読みを音訓辞書42より検索し、熟語における漢字の位置が左側でない場合には、語頭に濁音や半濁音があれば清音化し、一致した漢字の読みの部分を抽出する。読みの抽出が不可能な時、例えば、『煙草』(たばこ)のような通常の音訓に読みではない場合には、この熟語は使用しない。例えば、「垣」の場合、音訓辞書42に「カキ」があれば、先程の「カキ」「カ」のうち、「カキ」とそのまま一致することから、熟語の「垣根」が採用されることになる。

【0045】音訓辞書42に該当部分の語が存在し、自立語であるとき、その自立語の読み及びアクセント型を抽出する。自立語でない場合には熟語からの読み(連濁が発生していれば、その発生した情報の読み)を抽出し、アクセント型は“1型”を設定する。

【0046】次に、音韻情報生成部23は、熟語の読みと漢字の部分読みを助詞の「の」で、「熟語の読み」+「の」+「漢字の読み」の形式で接続し、「の」の直後にアクセント句境界を設定する。アクセント句毎のアクセント型の算出は、「NHK日本語発音アクセント辞典(平成3年11月10日第19刷)」の「共通語アクセント」の第2章「名詞のアクセント」、第6節「助詞の付いた形」に記載されているルールを用いて行うものとする。これにより、アクセント句毎の読みとアクセント

型が決定できる。

【0047】上記の一連の処理により、図9の2-Zの行に示すように、説明対象漢字『垣』、使用する熟語

「垣根」、生成される説明「垣根の垣」、生成される説明（アクセント付き仮名文）「カキ'ネノ/カキ'」が取得される。このように、取得された音韻情報及び韻律情報を音声合成装置3に入力することにより、音声合成装置3において音声合成処理を行い、合成音声を出力する。

【0048】〔第2の実施例〕以下に、第2の実施例を説明する。本実施例では、漢字入力部1から説明対象の漢字として『規』が入力された場合について説明する。

【0049】入力として「規」が漢字入力部1に入力されると、漢字入力部1は、当該漢字が手書き文字の場合には、文字認識処理を行い、認識結果として「規」と「項」を漢字候補として説明生成部2に転送する。この文字認識の結果は、利用者が「規」と入力したつもりであっても、手書き文字の場合、「規」に類似する漢字として、「項」という認識結果が得られたものとする。この「項」については後述する。

【0050】説明生成部2の熟語抽出部21は、『規』をキーとして熟語辞書41を検索する。図10に示す熟語辞書42を検索した結果、「規則」と「定規」が検索される。次に、「規」をキーにして図11に示す音訓辞書42を検索する。そこで、読みとして「キ」及び「ノリ」を抽出する。

【0051】また、位置確定部22は、「規則」の場合には、左側の文字、即ち、読みの前方の部分として、「キ」「キソ」が候補になど当該「規則」に対する位置を確定する。また、「定規」の場合は、右側の文字、即ち読みの後方の部分「ヨーギ」「ーギ」「ギ」が抽出されるが、拗音や長音で始まる単漢字は存在しないと判断し、ここでは、「ギ」のみを抽出する。右側、即ち読みの後方は、熟語で連濁、濁音化を生じる場合があることから、説明対象漢字「規」の読みであり、かつ清音化して一致するのは、どちらも「キ」であることから、この「ノリ」の読みは破棄される。促音や撥音で始まる語がないことも勿論利用できる。

【0052】さらに、説明フレーズ生成部23は、「規」の読みが「キ」のとき、これは、自立語でないため、熟語の読みを採用する。従って、「規則」では「キ」、「定規」では「ギ」が採用される。次に、説明フレーズ生成部23は、この読みを助詞の「ノ」で接続し、「規則のキ」、「定規のギ」を生成する。音韻情報生成部24は、このとき、アクセント句の境界は、「ノ」の直後とし、自立語でない部分の読みの部分「キ」と「ギ」には、どちらもアクセント型“1”を生成された説明フレーズに付与する。

【0053】ここで、音韻情報生成部24は、単語の読みと説明漢字部分の読みアクセントを付けて図9に示

すように、

「キ'ソクノ/キ'」「ジョ'ウギノ/ギ'」

のアクセント付き仮名文を生成する。

【0054】音韻情報生成部24で生成されたアクセント付き仮名文は音声合成装置3に転送され、音声合成装置3は、音声合成処理を行い、合成音声を出力する。また、図9において、(2-A)に示す漢字「規」と

「垣」に対して、使用する熟語が図9の(2-B)である時、生成される説明を図9の(2-C)、生成されるアクセント付き仮名文を図9の(2-D)に示す。

「垣」の字を説明する場合、「垣根」は熟語辞書41から検索され、「垣」は読みが「カキ」、アクセント型は“2”で、自立語（名詞）に登録されているため、「カキ'ネノ/カキ'」が生成される。

【0055】上記の実施例により、「カキネノカキ」で『垣』を、「キソクノキ」や「ジョウギノギ」で『規』を説明することが可能となる。なお、読みが異なっても清音化した読みが同じ場合を読みが類似した熟語と定義すると「定規（ジョーギ）」の読みと類似した「蒸気（ジョーキ）」が熟語辞書41に存在することから、熟語抽出部21では、熟語の「定規」を使用せずに、「規則」の方を採用するようにしてもよい。

【0056】また、漢字の手書き文字情報が入力された場合、漢字の候補と説明をペアにして出力することも可能であり、例えば、漢字入力部1における認識けつあととして上記のように『規』と『項』が取得された場合には、「1番は規則のキ1、2番は項目のコー」のように読み上げることで、正しい漢字候補を選択することも可能である。

【0057】音声合成部3は、アクセント付き仮名文を入力とする音声合成装置を用いることにより、音声を合成することができる。なお、用語辞書4の熟語辞書41は、利用者に利用傾向により熟語が登録されている辞書を用いるようにしてもよい。例えば、人名等を多く扱う場合には、人名に多く用いられる漢字を含む熟語を熟語辞書41に登録しておき、当該漢字の説明を熟語の読みと当該漢字の読みを説明フレーズとして音声出力することにより、電話番号案内や、電話オーダの受け付け等に適用することが可能である。また、あるジャンルにおける専門用語をよく用いる場合には、当該専門用語で構成される熟語を熟語辞書41に登録しておくことにより、当該ジャンルを扱う者にとっては容易に漢字を認識することができる。

【0058】また、熟語辞書41に登録される語彙として、良い印象または中立の印象の語、同音語が少ない語、良く知られている語を登録しておくことが望ましい。熟語を用語ととらえ、「渡」や「橋」をキーにして「渡る橋」で「ワタル」や「ハシ」の漢字を説明することもできる他、広→「広い」で、広（ヒロシ）は、『「広い」です』のように、説明することも可能であ

る。

【0059】〔第3の実施例〕次に、漢字に対する説明を漢字が有する旁等により説明する例について述べる。図12は、本発明の第3の実施例の漢字形状辞書の構成例を示す。同図において、漢字形状辞書8は、単漢字と、当該単漢字に対応する説明情報、構成情報、及び音韻検索番号から構成される。説明情報は、当該単漢字が含まれる熟語例、旁の説明、読み、当該漢字の種類を表す情報等により構成され、説明文を作成する際に参照される。構成情報は、単漢字の旁等のパーツ情報と当該パーツ情報の位置情報より構成され、同図の例では、例えば、『翔』には、パーツ情報「羊」と「羽」とがあり、「羊」の右側(R)に「羽」があることを示している。位置情報は、右側にパーツ情報が存在する場合には、(R)、下側にパーツ情報が存在する場合には(U)、パーツ情報の中にパーツ情報が位置する場合には(I)等が記入されている。

【0060】また、音韻検索番号は、音韻辞書9を検索する際の呼び出し番号であり、説明フレーズは6は、当該音韻検索番号に基づいて音韻辞書を検索する。図13は、本発明の第3の実施例の音韻辞書の構成を示す。同図に示す音韻辞書9は、音韻番号と音韻情報から構成され、それぞれの音韻番号に対応してカタカナで漢字の旁を説明している。音韻情報の“/”は、区切り記号である。この記号に沿って、音声合成部7は音声合成する。区切り記号を用いることにより、特に、同じ属性を列挙する際に、通常の読み上げでは、アクセント結合が生じると不自然になることから並列としてのアクセント、ポーズを付与する意味もある。例えば、『千』については、“イチ/ジユウ/ヒャク/セン/ノ/セン”とする。

【0061】また、音韻情報において、人と人の対話では、旁と旁の接続表現として「に」が最も多いが、「という字に」「(を)書いて」「の下に」「の右に」等の表現も存在する。情報として曖昧さの削減が可能な「の下に」は、漢字形状辞書の位置情報で生成している。また、「山偏(へん)に…」と表現することで、「山の右…」と等価であることから、重複した説明の生成を抑制する。

【0062】例えば、『嶋』については、“ヤマヘンにトリ”、『鳥』については“ヤマの下にトリ”という区別が可能となる。また、要素を漢字に拡張した場合、『青島』を“青いに島”ではなく、“青い島”のような接続ができるように、接続表現は用言の連体形接続の形式をとるものとする。

【0063】また、以下に具体的な例を用いて説明する。説明対象漢字として『翔』が漢字入力部5から入力されたものとして説明する。これにより、説明フレーズ生成部6は、漢字形状辞書8を検索して、音韻検索番号「1」を取得する。次に、説明フレーズ生成部6は、音

韻番号をキーとして音韻辞書9を検索して、当該漢字

『翔』に対応する説明フレーズとして「ヒツジ/ノ/ミギガワ/ニ/ハネ」を取得する。取得した説明フレーズを音声合成部7に転送することにより、音声合成部7は、当該説明フレーズを音声合成処理して出力する。

【0064】上記の実施例では、漢字形状辞書8に音韻検索番号を付与した構成であるが、漢字形状辞書8中に音韻検索番号を用いずに、構成情報に基づいて音韻辞書9を検索するようにしてもよい。この場合は、音韻辞書9において、構成情報に対応する情報を保持し、当該情報から音韻情報が検索できるように構成する。

【0065】例えば、図14に示すように、旁に関する音韻情報と位置情報(R, U, I等)に対応する音韻情報を含めて音韻辞書9を構成する。このような音韻辞書9を利用する場合には、上記の例では、漢字形状辞書8から「羊R羽」という構成情報を取得すると、まず、『羊』を用いて音韻辞書9を検索し、「ヒツジ」を取得し、次に、『R』を用いて音韻辞書9を検索し、「ノ/ミギガワ/ニ」を取得し、さらに、『羽』を用いて音韻辞書9を検索して「ハネ」を取得して、それぞれの音韻情報を結合して、音声合成部8に転送する。

【0066】また、音韻辞書でR→「ニ」とおくことで、「ヒツジ/ニ/ハネ」と生成することも可能である。また、前述の第1、第2の実施例と本実施例を合成して利用することも可能である。例えば、熟語により第1の説明対象の漢字の説明フレーズを取得し、更に当該漢字に付与するようなものがある場合には、第3の実施例を用いて構成される漢字を説明することも可能である。

【0067】例えば、説明対象の漢字『繁』について、第1または第2の実施例により、『俊敏の敏』(シュンビンノビン)を取得し、さらに第3の実施例により、構成情報として“敏U糸”を設定することにより、生成される説明フレーズは“シュンビン/ノ/ビン/ニ/イト”となる。

【0068】なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【0069】

【発明の効果】上述のように、本発明の漢字説明方法及び装置によれば、説明対象の漢字を個々に説明する文を熟語等の辞書を用いることにより、自動的に生成でき、シカモ、該当する漢字と読み情報の自立語の有無を調べることにより正しい読み情報とアクセント情報が付与されているため、自然な合成音声の出力が可能となる。

【0070】また、説明対象の漢字に対する複数の説明を併記して説明することにより、漢字の説明に対して曖昧さを減らした説明が出力されるため、従来のように、アナウンサや新聞等の校正作業用のみならず、広範囲なメッセージの入力確認に利用できる。

【0071】また、音声による自動説明が可能のため、

電話による姓名や企業名の漢字の確認等、これまでなされていなかったデータベース検索やカスタマからの注文受け付け等への応用も可能である。また、ワードプロセッサに組み込むことにより、音声で漢字文字を確認する場合にも威力を発揮する。

【図面の簡単な説明】

- 【図 1】本発明の原理を説明するための図である。
 【図 2】本発明の原理構成図である。
 【図 3】本発明の漢字説明装置の構成図である。
 【図 4】本発明の説明生成部の構成図である。
 【図 5】本発明の漢字説明処理の一連の動作のフローチャートである。
 【図 6】本発明の説明生成部の処理のフローチャートである。
 【図 7】本発明の第 2 の漢字説明装置の構成図である。
 【図 8】本発明の第 2 の漢字説明処理の動作のフローチャートである。
 【図 9】本発明の実施例を説明するための図である。
 【図 10】本発明の実施例の熟語辞書の例を示す図である。
 【図 11】本発明の実施例の音訓辞書の例を示す図である。
 【図 12】本発明の第 3 の実施例の漢字形状辞書の構成

例を示す図である。

【図 13】本発明の第 3 の実施例の音韻辞書の構成例を示す図である。

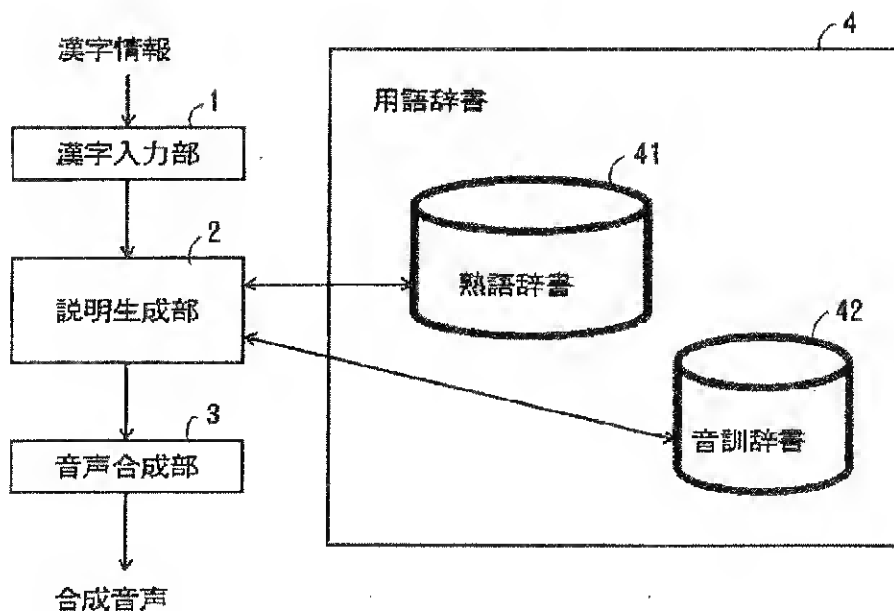
【図 14】本発明の第 3 の実施例の音韻辞書の他の構成例を示す図である。

【符号の説明】

- 1 漢字入力部
 2 説明生成部
 3 音声合成部
 4 用語辞書
 5 漢字入力部
 6 説明フレーズ生成部
 7 音声合成部
 8 漢字形状辞書
 9 音韻辞書
 10 漢字入力手段
 20 辞書
 30 熟語抽出手段
 40 位置確定手段
 41 熟語辞書
 42 音訓辞書
 50 説明文生成手段
 60 音声出力手段

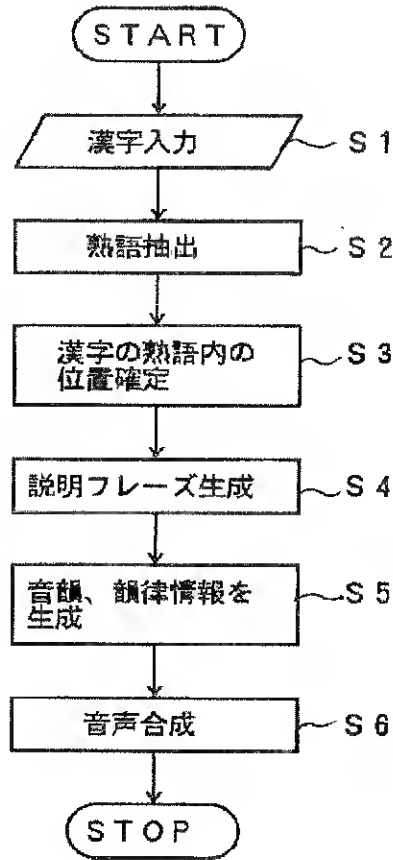
【図 3】

本発明の漢字説明装置の構成図



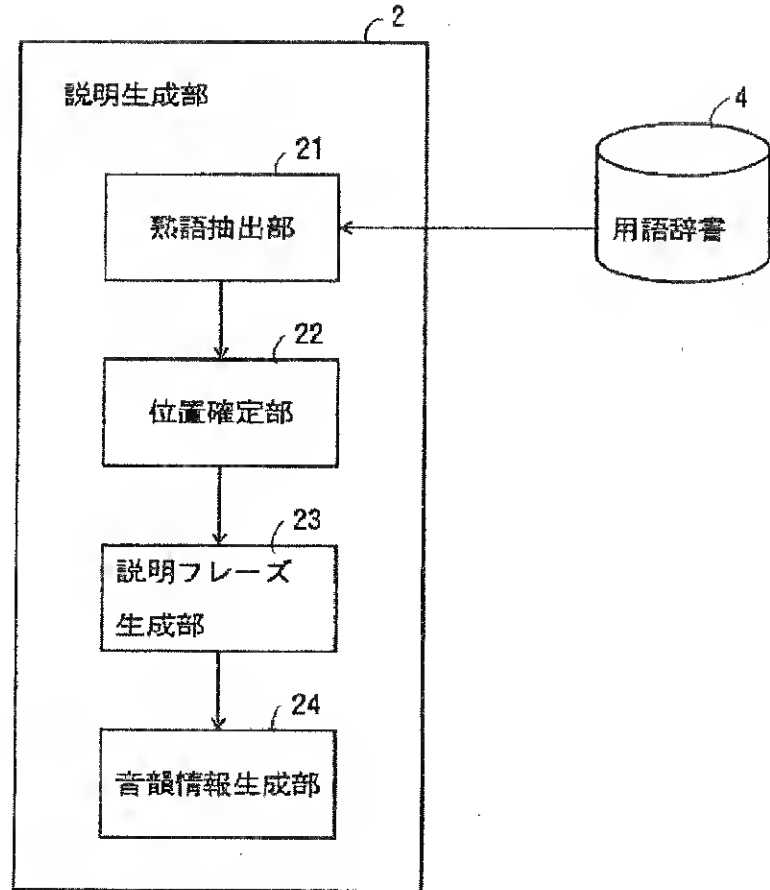
【図 1】

本発明の原理を説明するための図



【図 4】

本発明の説明生成部の構成図



音韻情報／韻律情報

【図 12】

本発明の第 3 の実施例の漢字形状辞書の機能例を示す図

漢 字	説明情報	構成情報	音韻検索番号
翔	飛翔	(羊 R 羽)	1
浩	サンズイー右一告	(#104 R 告)	2
杏	あんず／果実	(木 U 口)	3
国	ロー王	(口 I 王)	4

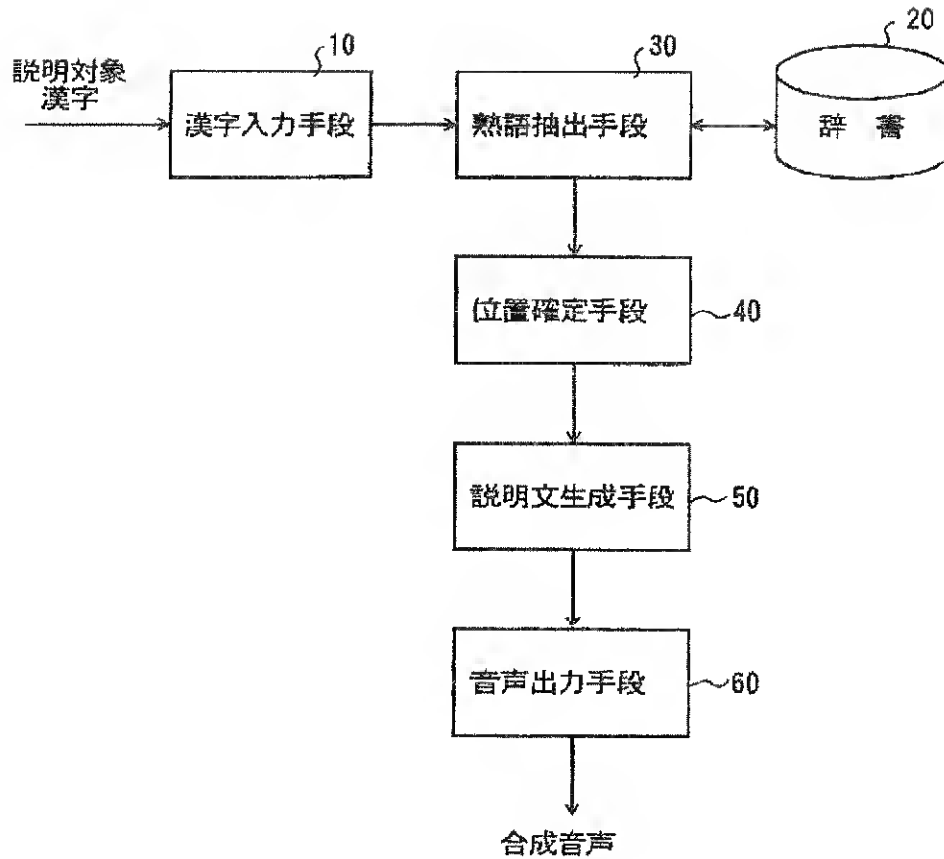
【図 13】

本発明の第 3 の実施例の音韻辞書の構成例を示す図

音韻番号	音韻情報
1	ヒツジ／ノ／ミギガワ／ニ／ハネ
2	サンズイ／ノ／ミギガワ／ニ／ツゲル
3	キ／ノ／シタ／ニ／クチ
4	クチ／ノ／ナカ／ニ／オウ

【図2】

本発明の原理構成図



【図11】

本発明の一実施例の言割辞書の例を示す図

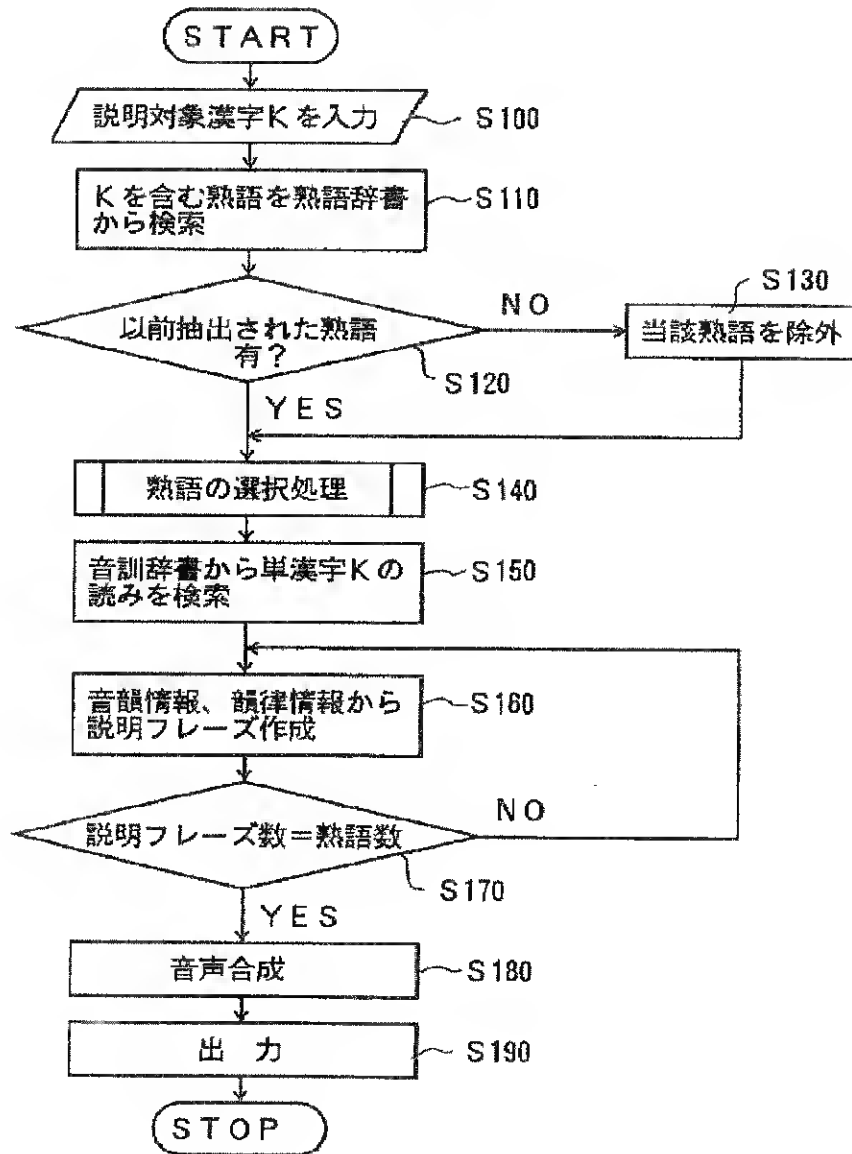
見出し&キー	読み	品詞	アクセント型
塩	カキ	名詞	2
規	ノリ	*	――
規	キ	*	――
項	コー	名詞	1
定	サダ・メル	動詞	3
定	ジョー	*	――

* : 接辞以外の造語成分

―― : 未収録

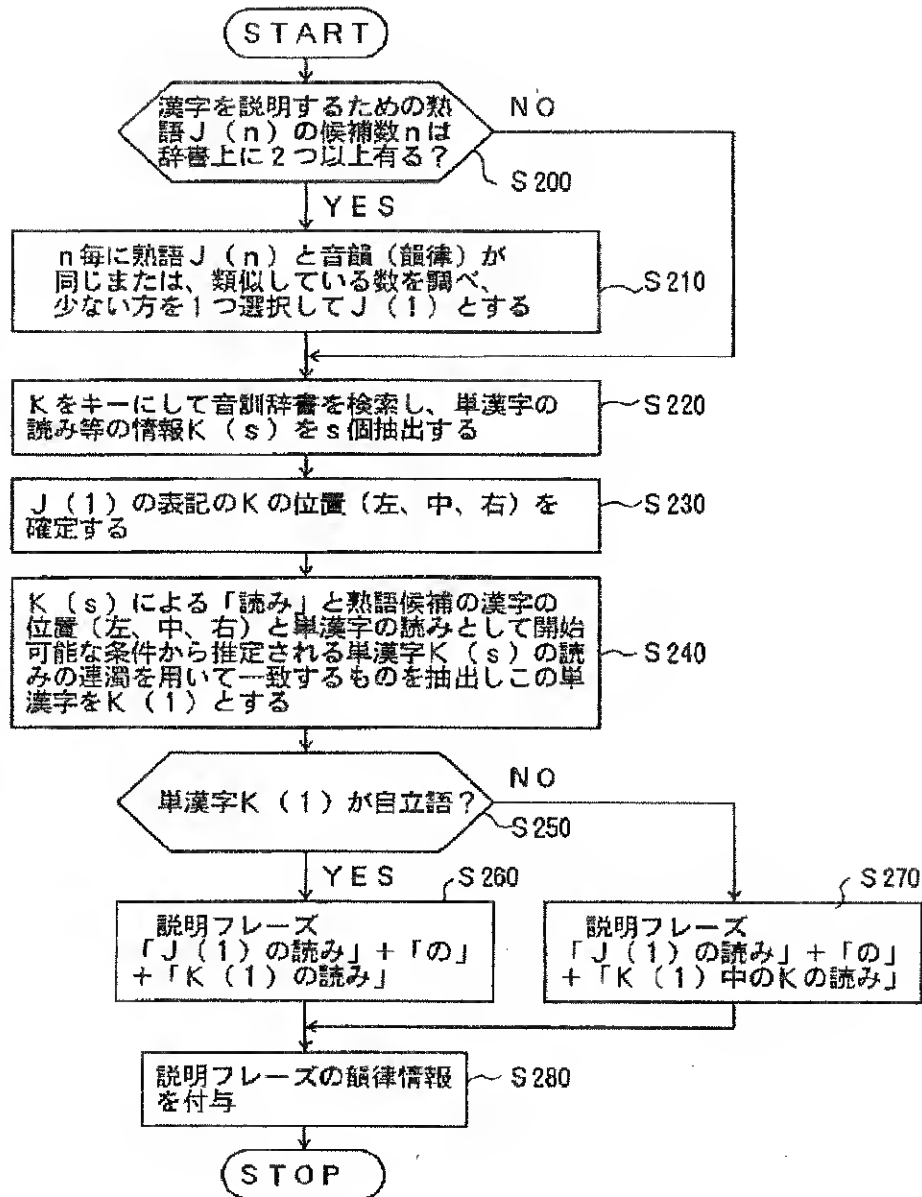
【図5】

本発明の漢字説明処理の一連の動作のフローチャート



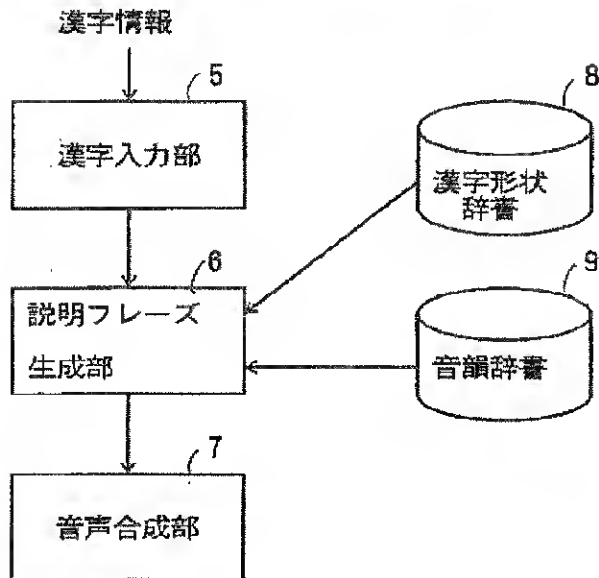
【図6】

本発明の説明生成部の処理のフローチャート



【図7】

本発明の第2の漢字説明装置の構成図



【図10】

本発明の一実施例の熟語辞書の例を示す図

見出し	読み	アクセント型
規則	キソク	1
垣根	カキネ	2
定規	ジョーギ	1
項目	コモク	0
蒸気	ジョーキ	1

【図9】

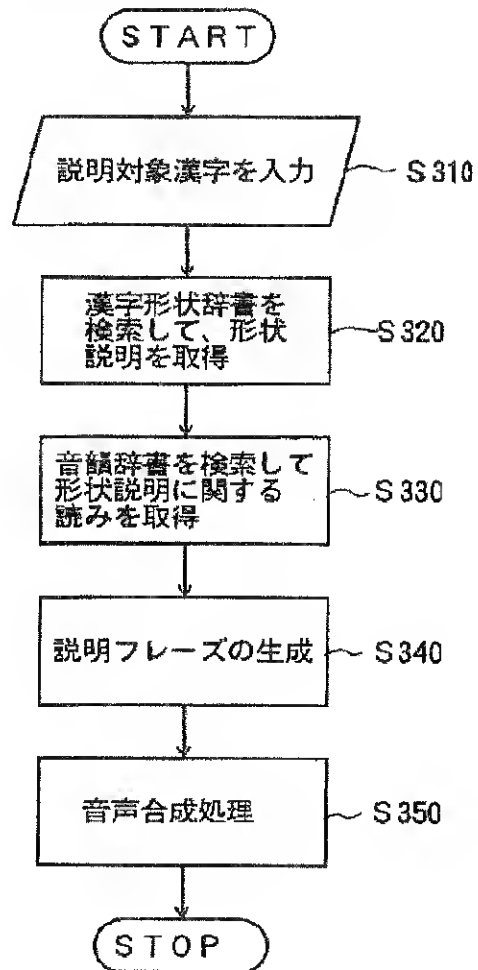
本発明の一実施例を説明するための図

	2-A	2-B	2-C	2-D
2-W	説明対象漢字	使用する熟語	生成される説明	生成される説明 (アクセント付) (仮名文)
2-X	規	規則	規則のキ	キ' ソクノ／キ'
2-Y	規	定規	定規のキ	ジョ' ウーギノ／キ'
2-Z	垣	垣根	垣根の垣	カキ' ネノ／カキ'

' : アクセントの下がり目を示す
/: アクセント句の境界を示す

【図8】

本発明の第2の漢字説明処理の動作のフローチャート



【図14】

本発明の第3の実施例の音韻辞書の他の構成例を示す図

パーツ情報	音韻情報
羊	ヒツジ
羽	ハネ
口	クチ
王	オウ
~	~
R	ノ/ミギガワ/ニ
U	ノ/シタ/ニ
I	ノ/ナカ/ニ
#104	サンズイ